Proaperiodic monoids via prime models

or: what do profinite words look like?

Sam van Gool

IRIF, Université Paris Cité

joint work with Benjamin Steinberg, City College of New York

Structure meets Power

ICALP workshop, Paris, 4 July 2022

Finite words and logic

Let Σ be a finite alphabet.

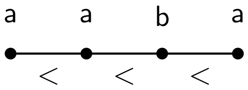
The set of finite words over Σ is a free **monoid** generated by Σ.

Finite words and logic

Let Σ be a finite alphabet.

The set of finite words over Σ is a free **monoid** generated by Σ.

A word $w \in \Sigma^*$ may be viewed as a finite structure with linear order $<$ and a decomposition into unary predicates $(P_a)_{a \in \Sigma}$:

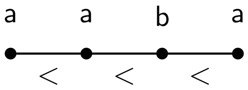


Finite words and logic

Let Σ be a finite alphabet.

The set of finite words over Σ is a free **monoid** generated by Σ.

A word $w \in \Sigma^*$ may be viewed as a finite structure with linear order $<$ and a decomposition into unary predicates $(P_a)_{a\in\Sigma}$:



Every **monadic second order sentence** in signature

$$S_\Sigma := \{<\} \cup \{P_a \ : \ a \in \Sigma\}$$

describes a set of finite Σ-words.

Regular sets

A set L of finite words is **regular** if it satisfies the following equivalent conditions:

- L is definable by a monadic second order sentence,

Regular sets

A set L of finite words is **regular** if it satisfies the following equivalent conditions:

- L is definable by a monadic second order sentence,
- L is recognizable by a finite automaton,

Regular sets

A set L of finite words is **regular** if it satisfies the following equivalent conditions:

- L is definable by a monadic second order sentence,
- L is recognizable by a finite automaton,
- L is saturated under a finite index monoid congruence on Σ^*,

Regular sets

A set L of finite words is **regular** if it satisfies the following equivalent conditions:

- L is definable by a monadic second order sentence,
- L is recognizable by a finite automaton,
- L is saturated under a finite index monoid congruence on Σ^*, i.e., there exists a surjective homomorphism

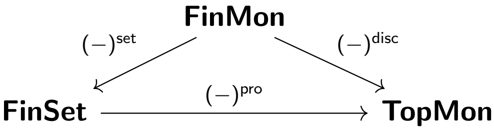
  $$h \colon \Sigma^* \twoheadrightarrow M,$$

  with M a finite monoid, such that, for some $P \subseteq M$,

  $$L = h^{-1}(P).$$

The free profinite monoid

The **free profinite monoid** over Σ is the, up to isomorphism unique, embedding of Σ into a topological monoid Σ^{pro} such that, for every finite monoid M and function $f\colon \Sigma \to M^{\mathsf{set}}$, there exists a unique continuous homomorphism $\overline{f}\colon \Sigma^{\mathsf{pro}} \to M^{\mathsf{disc}}$ that extends f.



$$\frac{f\colon \Sigma \to M^{\mathsf{set}}}{\overline{f}\colon \Sigma^{\mathsf{pro}} \to M^{\mathsf{disc}}}$$

Elements of Σ^{pro} are called **profinite words** over Σ.

Characterizing first-order logic

Theorem. A set $L \subseteq \Sigma^*$ is **first-order definable**

if, and only if,

Characterizing first-order logic

Theorem. A set $L \subseteq \Sigma^*$ is **first-order definable**

if, and only if,

L can be computed by an **aperiodic** finite monoid, i.e., one satisfying the profinite equation

$$x^{\omega} = x^{\omega} x.$$

Characterizing first-order logic

Theorem. A set $L \subseteq \Sigma^*$ is **first-order definable**

if, and only if,

L can be computed by an **aperiodic** finite monoid, i.e., one satisfying the profinite equation

$$x^{\omega} = x^{\omega} x.$$

Here, for any $x \in \Sigma^{\mathrm{pro}}$, the ω-power x^{ω} of x is defined as the unique idempotent element in the orbit-closure of x.

Characterizing first-order logic

Theorem. A set $L \subseteq \Sigma^*$ is **first-order definable**

if, and only if,

L can be computed by an **aperiodic** finite monoid, i.e., one satisfying the profinite equation

$$x^{\omega} = x^{\omega} x.$$

Here, for any $x \in \Sigma^{\mathrm{pro}}$, the ω-power x^{ω} of x is defined as the unique idempotent element in the orbit-closure of x.

Aperiodicity is equivalent to the absence of non-trivial subgroups.

We get the monoid of **proaperiodic words** as the quotient

$$\Sigma^{\mathrm{ap}} := \Sigma^{\mathrm{pro}} / (x^{\omega} = x^{\omega} x).$$

Schützenberger 1965; McNaughton & Papert 1971; Reiterman 1982

Constructions of the free profinite monoid

Theorem. The topological space underlying the free profinite monoid Σ^{pro} can be constructed as:

- the limit in **TopMon** of a projective diagram of finite monoids,
- an ultrametric completion of Σ^*,
- the ultrafilter space of the Boolean algebra of regular sets.

Constructions of the free profinite monoid

Theorem. The topological space underlying the free profinite monoid Σ^{pro} can be constructed as:

- the limit in **TopMon** of a projective diagram of finite monoids,
- an ultrametric completion of Σ^*,
- the ultrafilter space of the Boolean algebra of regular sets.

Theorem. The multiplication on Σ^{pro} is dual to a residuation structure on the regular subsets of Σ.

Reiterman 1982; Gehrke, Grigorieff & Pin 2008

Constructions of the free profinite monoid

The equivalence

$$\text{regular} \Leftrightarrow \text{MSO-definable}$$

induces a homeomorphism

$$\Sigma^{\text{pro}} \cong \text{completions of the MSO-theory of finite words.}$$

Constructions of the free profinite monoid

The equivalence

$$\text{regular} \Leftrightarrow \text{MSO-definable}$$

induces a homeomorphism

$$\Sigma^{\text{pro}} \cong \text{completions of the MSO-theory of finite words.}$$

Similarly, the equivalence

$$\text{aperiodic} \Leftrightarrow \text{FO-definable}$$

induces a homeomorphism

$$\Sigma^{\text{ap}} \cong \text{completions of the FO-theory of finite words.}$$

Steinberg & G. 2016; Linkhorn 2021

What do proaperiodic words look like? (I)

Definition. A first-order structure W in signature S_Σ is **pseudofinite** if it is a model of the common theory of finite words.

What do proaperiodic words look like? (I)

Definition. A first-order structure W in signature S_Σ is **pseudofinite** if it is a model of the common theory of finite words.

Theorem. There is an isomorphism of topological monoids

$$\Sigma^{\mathsf{ap}} \cong \text{elementary equivalence classes of pseudofinite } \Sigma\text{-words.}$$

Here, multiplication of pseudofinite words is done by concatenation of class representatives.

What do proaperiodic words look like? (I)

Definition. A first-order structure W in signature S_Σ is **pseudofinite** if it is a model of the common theory of finite words.

Theorem. There is an isomorphism of topological monoids

$$\Sigma^{\mathsf{ap}} \cong \text{elementary equivalence classes of pseudofinite } \Sigma\text{-words.}$$

Here, multiplication of pseudofinite words is done by concatenation of class representatives.

Proposition (model theory). Every elementary equivalence class of pseudofinite Σ-words contains an ω-**saturated** model.

What do proaperiodic words look like? (I)

Definition. A first-order structure W in signature S_Σ is **pseudofinite** if it is a model of the common theory of finite words.

Theorem. There is an isomorphism of topological monoids

$$\Sigma^{\mathsf{ap}} \cong \text{elementary equivalence classes of pseudofinite } \Sigma\text{-words.}$$

Here, multiplication of pseudofinite words is done by concatenation of class representatives.

Proposition (model theory). Every elementary equivalence class of pseudofinite Σ-words contains an ω-**saturated** model.

This can be used, e.g., to solve the word problem of $\langle \Sigma^{\mathsf{ap}}, \cdot, ()^{\omega} \rangle$.

Steinberg & G. 2016

Example (I)

The saturated model representing the proaperiodic word $(ab)^\omega$ is

$$ababab\ldots \quad (\ldots ababab\ldots)^{\mathbb{Q}} \quad \ldots ababab,$$

where the middle part has order type $\mathbb{Q} \times_{\mathrm{lex}} \mathbb{Z}$.

Example (I)

The saturated model representing the proaperiodic word $(ab)^\omega$ is

$$ababab\ldots \quad (\ldots ababab\ldots)^{\mathbb{Q}} \quad \ldots ababab,$$

where the middle part has order type $\mathbb{Q} \times_{\mathrm{lex}} \mathbb{Z}$.

A calculus: finite structures are saturated, substitution of saturated models into saturated models remains saturated.

Example (I)

The saturated model representing the proaperiodic word $(ab)^\omega$ is

$$ababab\ldots \quad (\ldots ababab\ldots)^{\mathbb{Q}} \quad \ldots ababab,$$

where the middle part has order type $\mathbb{Q} \times_{\text{lex}} \mathbb{Z}$.

A calculus: finite structures are saturated, substitution of saturated models into saturated models remains saturated.

There also exist proaperiodic words with uncountably many factors (these are a bit harder to draw).

Proaperiodic words and step points

An alternative way of "realizing" profinite words as structures:

Definition. For a proaperiodic word u, define a **point** of u as a pair $(u_1, u_2) \in (\Sigma^{\mathsf{ap}})^2$ such that $u_1 u_2 = u$.

Proaperiodic words and step points

An alternative way of "realizing" profinite words as structures:

Definition. For a proaperiodic word u, define a **point** of u as a pair $(u_1, u_2) \in (\Sigma^{\mathsf{ap}})^2$ such that $u_1 u_2 = u$.

The set of points of u may be linearly ordered, in such a way that the following theorem holds:

Theorem. A proaperiodic word can be fully described by a labeled linear order of its **step points**, i.e., those points that have an immediate predecessor and successor, or none at all.

Almeida, A. Costa, J. Costa, Zeitoun 2017

What do proaperiodic words look like? (II)

The step point structure fits well with our point of view.

Recall that a model of a theory T is **prime** if it embeds elementarily into every model of T.

What do proaperiodic words look like? (II)

The step point structure fits well with our point of view.

Recall that a model of a theory T is **prime** if it embeds elementarily into every model of T.

Theorem (model theory + some work). Every elementary class of pseudofinite words contains a **prime** model, which is isomorphic to the step point structure (up to an off-by-one error), and multiplication of step point structures is just concatenation of prime models.

Steinberg & G. 202x

Example (II)

The prime model representing the proaperiodic word $(ab)^\omega$ is

$$ababab\ldots \qquad\qquad \ldots ababab,$$

where the middle part has disappeared.

What do profinite words look like?

Some possible further directions:

- What happens for MSO on more than one letter?
- What happens for fragments of FO (in particular $B\Sigma_n$)?
- What happens for profinite structures other than words?
- What is the correct categorical point of view on all of this?